AN INTRODUCTION TO SHAPE DYNAMICS

PATRICK KERRIGAN

Abstract. Shape Dynamics (SD) is a new fundamental framework of physics which seeks to remove any non-relational notions from its methodology. importantly it does away with a background space-time and replaces it with a conceptual framework meant to reflect direct observables and recognize how measurements are taken. It is a theory of pure relationalism, and is based on different first principles then General Relativity (GR). This paper investigates how SD assertions affect dynamics of the three body problem, then outlines the shape reduction framework in a general setting.

1. Introduction

Shape Dynamics was introduced by Julian Barbour. Subsequent work on best matching as SD formulation of N-body Newtonian gravity was done in collaboration with Bruno Bertotti. More recent work had been done with Flavio Mercati and Tom Koslowski on gravity, entropy and the arrow of time. Henrique Gomes discusses the principle bundle structure in Geometrodynamics, which is the shape dynamic analog of GR. Many others have also published papers on the subject.

This paper will first cover a toy model, the shape dynamic 3-body problem. We reduce the space of configurations by identifying all similar triangular states the 3 particles could take. Barbour and others refer to this as 'removing space'. Though the concept of background space is obviously useful and fundamental for essentially all of modern physics, we do not ever observe space and time directly. Rather we arrive at their abstraction from observable relations among particles alone. According to Barbour, this is called *Mach's principle* by Einstein[1], named after physicist Ernst Mach, as he appears to be the first to suggest and seriously consider its implications in *The Science of Mechanics*[7]. Poincaré did work related to this problem independently, and came to the conclusion that determining the trajectory of a physical system from observable relational information alone is impossible. This is because, as Poincaré showed, one cannot deduce the angular momentum of the system from such information only[10]. Shape Dynamics (SD), and its Newtonian gravity analog in Barbour-Bertotti best maching, shows that a theory with no such ambient space, but only relational data, can reduce to locally defined notions of distance and time that are equivalent to the classical understanding, as well as determining physical trajectories from purely relational data. These come from "Clocks" and "Rods", which is matter 'marching in step' in the case of clocks, and matter with extent collected together, defining "Rods". Marching in step simply means there there is some agreed upon method for tracking the changes of things, and can be used to keep appointments. Put simply, one "fixes" pure relational trajectories by setting the "missing" information, (specifically the angular momentum) to zero.

Essential to shape dynamics is the assertion that the meaning of *extent* in both the spacial and temporal sense are not related to some background or ambient thing in which the object has extent, eg. space and time, but rather to other objects that also have similar extent nearby. To measure something with a meter stick there are two things present: the meter stick and the thing being measured. Not some ideal 'meter' which is an intrinsic property of the ambient space itself.

Objects in the universe have extent. The program of shape dynamics is to remove absolute notions of size, and instead asserts that all thing exist only in relation to other things; it is fundamentally an assertion that purely relational data is the only physically real data. This assertion, called by Mercati[8] *spacial relationalism*, has consequences on the variety of dynamical phenomena available to the theory, and the mathematical framework describing these dynamical restriction is recognizably a *gauge theory*. Gauge theories have a mathematical structure called a *principle bundle*, which is one way of mathematically formalizing their features. This paper uses some category theory, steaming from the desire for a pictorial representation of the logical arguments involved. Category theory is a powerful, sophisticated, and nuanced mathematical program, but we use only some basic notions here and rely mostly on its intuitive feel and pictorial abilities.

The theory of principle bundles is a powerful tool for discussing any gauge theory.[1] Central to the study of these mathematical objects is group theory, particularly Lie groups. Also discussed are some general properties of manifolds and conservation laws from Noethers' theorem specifically in the context of SD. Hamiltonian and Lagrangian mechanics in the language of category theory will tie all of these ideas together in a simple pictorial representation for the discussion of the logical relations between them. This has the benefit of being clean to look at and understand in the abstract, but hard to ground in direct calculation techniques, so some arithmetic simplifications for some of the related concepts will also be discussed.

2. TOY MODEL: THE 3-BODY PROBLEM WITH SHAPE RESTRICTIONS

Lets start with a toy model 3 particle Newtonian universe and investigate how it acts once shape dynamical restrictions are applied. Specifically we re-build the 3 particle shape space discussed by Barbour et. al. [3], called the shape sphere. Mathematically this is a moduli space of triangles, though perhaps triangles is an improper description, as our space include co-linear configurations which are not quite triangles in the classical geometric sense—they have no area in their interior—as well as singular binary collision points. Our example space also tracks orientation, since the "shape sphere" includes a 2-fold, 6-fold if we forget particle labels/masses, degeneracy of shape representatives to maintain a smooth dynamical trajectory of the flow generated by Hamilton's equations. In a mathematical sense this means the similarity transformation torsors intersect the space sphere twice each.

2.1. the 3-body problem.

Consider the classical Newtonian gravitational 3 body problem. Given some initial conditions $(\vec{x}_i, \vec{p}_i) \in \mathbb{R}^9 \times T^*\mathbb{R}^9 \simeq \mathbb{R}^{18}$. Here we should note that $\vec{p}_i$ is more mathematically rigorously described as a covector in the cotangent space of $\mathbb{R}^9$, but because our coordinates are an $\mathbb{R}$-vector space, the cotangent space is just another isomorphic copy of $\mathbb{R}^9$. This distinction becomes important in the GR extension. Additionally it might be even more accurate to describe $\mathbb{R}^9$ as $(\mathbb{R}^3)^3$, since the geometric object being described is all possible triples of $\mathbb{R}^3$-vectors. This is why we here use the notations of vector hats on our particle labels indexed by $i \in 1, 2, 3$ a particle label set. With particle masses m_i, the dynamics is governed by the Hamiltonian,

$$\mathcal{H} := \sum_i \frac{\vec{p}_i \cdot \vec{p}_i}{2m_i} - \sum_{i,j<i} \frac{m_i m_j}{r_{ij}} \tag{1}$$

where $r_{ij} := |\vec{x}_i - \vec{x}_j|$ and $G := 1$, which we can do by suitable choice of units without loss of generality. Hamilton's equations specify a dynamical trajectory,

$$\begin{aligned} \vec{x}_i :& \mathbb{R} \to \mathbb{R}^9 \\ & t \mapsto (\vec{x}_1(t), \vec{x}_2(t), \vec{x}_3(t)) \end{aligned} \tag{2}$$

such that

$$\begin{aligned} \frac{\partial \vec{p}_i}{\partial t} &= -\frac{\partial \mathcal{H}}{\partial \vec{x}_i} \\ \frac{\partial \vec{x}_i}{\partial t} &= +\frac{\partial \mathcal{H}}{\partial \vec{p}_i} \end{aligned} \tag{3}$$

Where again in this case $\vec{p}_i = m_i \dot{\vec{x}}_i$.

This differential equation is clearly non-linear (becuase of the $\frac{1}{r_{ij}}$) and therefor analytic solutions do not exist in every case, but every initial condition defines a unique trajectory, via Hamilton's equations, that is computationally solvable to arbitrary accuracy. One can show that with the data $\vec{x}_i(0), \vec{p}_i(0)$, one gets a well defined trajectory in the coordinate space $\mathbb{R}^9$ which is the entire history of that 3 particle Newtonian universe. The fundamental argument of shape dynamics is this: in a 'real' 3 particle universe, the only data that exist is some set of shape descriptive numbers—like the ratios of side lengths of the triangle formed by the 3 particles, r_{12}/r_{23} and r_{13}/r_{23}, and their conjugate momenta, where the choice of denominator is arbitrary—and *no other information.* This is not the only choice we could have made. Any set of numbers which describe the shape data only would work, for example, two of the interior angles formed by the triangle.

[1]Mercati in [8] suggests that the first to recognize the principle bundle structure of SD was H. Gomes

Another way to say this mathematically is that "every similar triangle formed by the 3 particles is physically indistinguishable and identical". It is now useful to try and mathematically classify all the distinguishable configurations as this is the 'real space' of observable configuration.

Take a moment to recognize how abstract this concept really is, as there is no way to consider a 'person' or observer in a 3 particle universe, you can't make a person with 3 particles! Clearly this is beyond the realm of verification as far is physical theories go, so what were really trying to build here is a utilitous metaphor for more complete physics, the ultimate goal being a unified theory. Such a theory is well beyond the scope of this paper, I hope to cover only a simple model to help lift to more nuanced models.

2.2. **3-particle shape reduction.** The reason we construct this moduli space of triangles, called the shape space, is to try and solve dynamical problems in it. We feel that every initial condition in this shape space should lead to the same dynamics even though they could have come from different, but shape equivalent, initial Newtonian conditions. This is from the postulates of shape dynamics, that similar Newtonian dynamics are equivalent shape dynamics. 'Shape equivalent trajectory' means that the trajectory of the system in shape space is exactly the same for two given initial conditions in the original Newtonian absolute space.

Lets assume we have 3 particle trajectories satisfying Newton's gravity equations in $\mathbb{R}^3$, labelled $\vec{x}_1(t)$, $\vec{x}_2(t)$ and $\vec{x}_3(t)$. At each instant in time these three vector-functions are together determined by 9 real numbers, and define a triangle of points in $\mathbb{R}^3$. We reduce the dynamics following the process used by Barbour et. al. [3] and others (additional citations). First we insist that the origin be the center of mass at all times. This is a result of the translation reduction and the choice of origin as center of mass is from trying to make a translation reduced coordinate choice. This has a result of fixing the total translational momentum of the universe as zero, since the center of mass has no velocity for all time. Mathematically we perform a linear transformation to Jacobi coordinates $\vec{\rho}_i$ using a matrix of masses,

$$\vec{\rho}_b = M^a_b \vec{x}_a \tag{4}$$

with,

$$M^a_b = \begin{bmatrix} -\sqrt{\frac{m_1 m_2}{m_1+m_2}} & \sqrt{\frac{m_1 m_2}{m_1+m_2}} & 0 \\ -\sqrt{\frac{m_3}{(m_1+m_2+m_3)(m_1+m_2)}}m_1 & -\sqrt{\frac{m_3}{(m_1+m_2+m_3)(m_1+m_2)}}m_2 & \sqrt{\frac{m_3(m_1+m_2)}{m_1+m_2+m_3}} \\ \frac{m_1}{\sqrt{m_1+m_2+m_3}} & \frac{m_2}{\sqrt{m_1+m_2+m_3}} & \frac{m_3}{\sqrt{m_1+m_2+m_3}} \end{bmatrix} \tag{5}$$

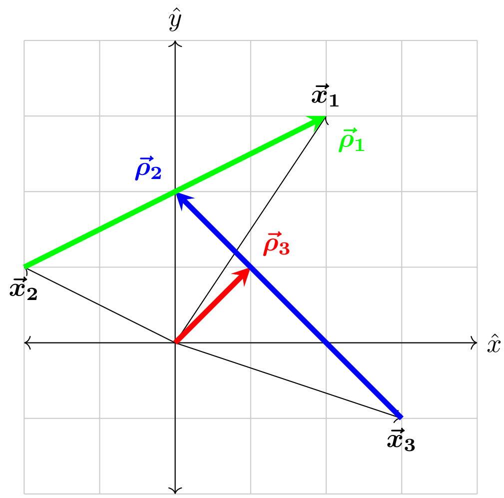
The $\vec{\rho}_3$ coordinate is the center of mass and so moving on down to triangle space, we 'forget', or omit by setting to zero, that coordinate, removing translations from the state description. Equivalently, we could say that the third row of the transformation should be all zeroes, or even that it doesn't exist at all, and that the transformation is carried out by a 2 by 3 matrix. These are all the same physical assertion: the the center of mass coordinate is irrelevant to the physics of the system. This is not a new assertion of shape dynamics, and has been known since the time of Newton that such a choice of origin can be made, it is called *Galilean invariance.*

It is important to recognize that the remaining state vectors $\vec{\rho}_1, \vec{\rho}_2$ are invariant under overall translations of the state. The $\vec{\rho}_3$ vector contained all of the 'overall' translation information, so by forgetting that coordinate, we constructed a space invariant under overall translations. Specifically, invariance under overall translation means that given an instantaneous state $(\vec{x}_1, \vec{x}_2, \vec{x}_3)$, translated by an arbitrary 3 vector $\vec{b} \in \mathbb{R}^3$, to $(\vec{x}_1+\vec{b}, \vec{x}_2+\vec{b}, \vec{x}_3+\vec{b})$, and perform the same Jacobi coordinate transformation, giving the exact same $(\vec{\rho}_1, \vec{\rho}_2)$. Only $\vec{\rho}_3$ changes, to $\vec{\rho}_3 + \vec{b}$. By setting this coordinate to zero we eliminate overall translational dependence from our representation of the instantaneous physical state. Note that by ignoring the $\vec{\rho}_3$ coordinate, the number of real numbers we need to describe a state is reduced by 3—instantaneous configurations are now described by 6 independent real numbers rather then 9.

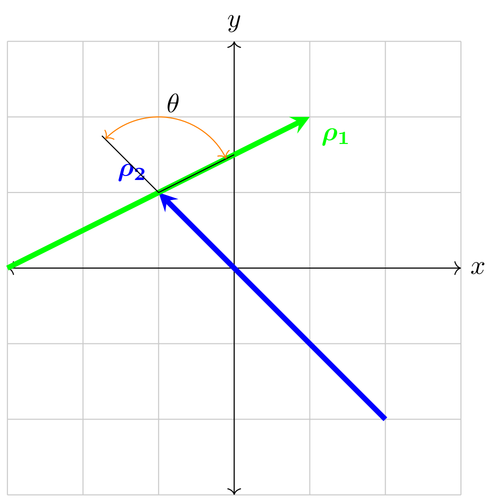
This transformation also acts on the momenta of the Cartesian coordinate trajectories to give Jacobi momenta,

$$\vec{\kappa}^b = M^{-1b}_a \vec{p}^a \tag{6}$$

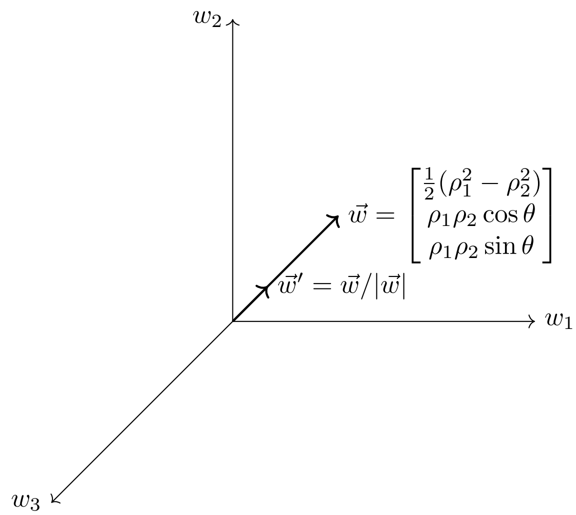
Note that if the third row is set to zero, M is non-invertible, however this is essentially a point about information being lost, specifically the center of mass coordinate. When the center of mass is set as the origin, the velocity, and thus also the momentum, of the center of mass of this universe is also set to zero ($\vec{\rho}_3 = 0 \implies \dot{\vec{\rho}}_3 = 0$) After this reduction was asserted, the momentum associated to the generalized



(A) 3 particles with coordinates $\vec{x}_i$ and Jacobi coordinates $\vec{\rho}_i$



(B) Detranslated Jacobi coordinates with center of mass 'forgotten' or set to zero



(C) w-coordinates, note $|\vec{w}'| = 1$

FIGURE 1. The shape reduction process on a sample 3 particle configuration

coordinate which co-varied with global translations was set to zero, as a consequence of the new coordinate being constant for all time. This is a complicated explanation of this simple example which will help with later similar arguments.

Next we perform another coordinate transformation, but this time non-linear, to what is called by Barbour et. al.[3] the w-coordinates.

$$w_1 = \frac{1}{2}(\vec{\rho}_1 \cdot \vec{\rho}_1 - \vec{\rho}_2 \cdot \vec{\rho}_2) = \frac{1}{2}(\rho_1^2 - \rho_2^2) \tag{7}$$

$$w_2 = \vec{\rho}_1 \cdot \vec{\rho}_2 = \rho_1\rho_2\cos\theta \tag{8}$$

$$w_3 = \vec{n} \cdot \vec{\rho}_1 \times \vec{\rho}_2 = \rho_1\rho_2\sin\theta \tag{9}$$

$\vec{n}$ is a vector chosen normal to the plane in which the particles stay for all time. These coordinates are invariant under the action of the rotations of 3 space, so this transformation forgets that information.

$$\begin{gathered}\forall \boldsymbol{R} \in SO(3), \\ \vec{\rho}_a \cdot \vec{\rho}_b = \boldsymbol{R}\vec{\rho}_a \cdot \boldsymbol{R}\vec{\rho}_b \\ \|\vec{\rho}_a \times \vec{\rho}_b\| = \|\boldsymbol{R}\vec{\rho}_a \times \boldsymbol{R}\vec{\rho}_b\|\end{gathered} \tag{10}$$

To show dot product invariance simply use the angle-magnitude representation

$$\vec{\rho}_a \cdot \vec{\rho}_b = \|\vec{\rho}_a\|\|\vec{\rho}_b\| \cos\theta \tag{11}$$

then by multiplicativity of vector norms with determinants of their matrix operators

$$\boldsymbol{R}\vec{\rho}_a \cdot \boldsymbol{R}\vec{\rho}_b = \|\boldsymbol{R}\vec{\rho}_a\|\|\boldsymbol{R}\vec{\rho}_b\| \cos\theta \tag{12}$$

$$= (\det \boldsymbol{R})^2 \|\vec{\rho}_a\|\|\vec{\rho}_b\| \cos\theta \tag{13}$$

$$= \vec{\rho}_a \cdot \vec{\rho}_b \tag{14}$$

This exact same argument works for the cross product term in the w_3 coordinate, because we simply substitute $\sin\theta$ for $\cos\theta$ and the rest of the argument doesn't change. Note again that we have moved from a 6 dimensional space of $(\vec{\rho}_1, \vec{\rho}_2)$, to the 3D w-coordinates by removing the 3D group of rotations dependence from our description of instantaneous configurations.

One can see how these w-coordinates still contain information about the shape of the triangle being described, as the angle formed by the w-vector projected in the w_2-w_3-plane measured to the w_2-axis is precisely the same angle θ, shown in Figure 1.

Part of forgetting overall rotations is setting the total center of mass angular momentum of the system to zero for all time, which is why the 3 particles always stay in the same plane. Finally, upon noting

$$\|\vec{w}\| = \rho_1^2 + \rho_2^2 = I_{CM}/2 \tag{15}$$

we can fix the center of mass inertia as 2 for all t, effectively fixing a choice of scale for the system, but restricting the 3-d w-coordinate space too the unit sphere by annihilating radial fibers,

$$\vec{w}' = \vec{w}/\|\vec{w}\| \tag{16}$$

These w'-coordinates are independent of overall changes of scale. Suppose the Jacobi coordinate undergo a transformation $(\vec{\rho}_1, \vec{\rho}_2) \to (s\vec{\rho}_1, s\vec{\rho}_2)$, for some arbitrary $s \in \mathbb{R}^+$ then the w-coordinates would change too $\vec{w} \to s^2\vec{w}$. However upon reducing to $\vec{w'} = s^2\vec{w}/\|s^2\vec{w}\| = \vec{w}/\|\vec{w}\|$, it is seen that w'-coordinates are independent of arbitrary changes of scale.

Thus we remove all the similarity transformations for the space. We are left with a unit sphere in the 3-D $\vec{w}$-coordinates. Each point of this sphere corresponds to a equivalence class of triangles, and so the initial trajectory is embedded into a single path on this sphere describing the shape dynamics of the system.

$$\vec{w}' : \mathbb{R} \to S^2 < \mathbb{R}^3 \tag{17}$$

$$t \mapsto \vec{w}'(t) \tag{18}$$

A picture of a sample trajectory is shown in Figure 2. This figure shows the shape sphere, with shape potential shown with coloring, and its gradient, the shape force, drawn with arrows on the right. On the left is the original Newtonian trajectory of the 3 particles. This trajectory was not solved in the shape space, but rather in the original coordinate space, with several of the shape restrictions imposed.

It is fairly easy to observe some of the key features of the shape restrictions. First the system falls into potential wells centered on binary collisions. These are the spirals moving towards the low potential red zones. The topology of the space and potential function means the system is guaranteed to move through a point of maximum potential, when it is closest to either of the blue north and south poles. Without going into more detail, there is reason to suspect that these points are a kind of minimum gravitational entropy point, and perhaps their existence in SD represent an explanation for the big bang. See [3],[5] for more details.

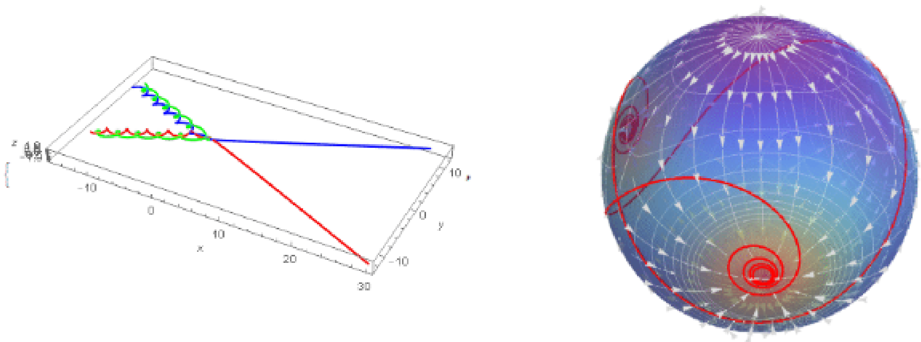


FIGURE 2. A Newtonian 3 body trajectory in Carteisan coordinates on the right, with shape conditions ($\vec{P}_{cm} := \sum_i \vec{p}_i = 0, \vec{L}_{cm} := \sum_i \vec{r}^i \times \vec{p}_i = 0, E = 0$ imposed, and the w'-coordinate shape sphere with corresponding projected trajectory on the right

An important note: if our argument were to consistently make sense, the *dilatation momentum* D of the system would be zero for all time. It is defined as,

$$D = \frac{\dot{I}_{cm}}{2} = \sum_i \vec{x}_i \cdot \vec{p}_i \tag{19}$$

In center of mass coordinates $\vec{x}_i$. This corresponds to the fact that we cannot 'leave' the sphere of shapes. We have fixed I_{CM} as constant, by restricting to w vectors of unit length. However, dilation's are not a symmetry of the given Newtonian potential, and so dilatation momentum is not conserved in such a system, due to Noether's theorem. If the assertions of shape dynamics are true, gravity *cannot obey a* $1/r_{ij}$ *potential.* This would seem to be a serious problem, we will discuss possible resolutions next.

2.3. **Triangle dynamics.** In order to realize the Mach-Poincaré principle, we need a dynamical trajectory for 3 particles determined entirely by their relational data, for 3 particles, the 3 r_{ij} and their first derivatives. In fact this contains a scale redundancy, so we may instead choose two pairwise ratios to eliminate this redundancy. We can always project any 3 particle configuration to the shape sphere using the reduction described above, but shape dynamics is not simply tracing the shape trajectory in the sphere. The assertion is that the sphere is the only data that exists, that really we ought not be able to leave it. We are saying that overall translation, rotations, and changes of scale *do not exist in physical reality.* This has the effect of fixing as constant 7 generalized coordinates in the Newtonian framework. In the triangle example, 3 coordinates describing the center of mass, $\vec{\rho_3} = 0$, 3 coordinates describing the angular state, which were unnamed, and 1 the center of mass moment of inertia, characterizing the scale, $I_{CM} = 2$. This means the momentum associated to these coordinates must be zero for all time[3]. These are

$$\vec{P}_{tot} = \sum_i \vec{p}_i = 0 \tag{20}$$

$$\vec{L}_{tot} = \sum_i \vec{x}_i \times \vec{p}_i = 0 \tag{21}$$

$$D = \sum_i \vec{x}_i \cdot \vec{p}_i = 0 \tag{22}$$

Because translations and rotations are symmetries of the given Newtonian Hamiltonian, we can just subtract any initial translational and rotational momentum for some initial conditions, and know that they will be conserved at zero for all time. However changes of scale are not a symmetry of the given system. We cannot impose the third restriction for classical n-body gravity. D is not a conserved quantity ($\dot{D} \neq 0$), which contradicts it being zero for all time ($D = 0 \implies \dot{D} = 0$). We can fix this problem by replacing the Newtonian potential with one that produces a scale conserving Hamiltonian

$$U_{new} \to U := \frac{U_{new}}{\sqrt{I_{cm}}} \tag{23}$$

We make this change because a given potential which is homogeneous of degree k, it can be shown from Newtons second law that

$$\ddot{I}_{cm} = 2\dot{D} = 4E - 2(k+2)U \tag{24}$$

By fixing $E = 0$ and $k = -2$, we get $\dot{D} = 0$ as we would hope. This is an interesting step, because to get a conserved dilatational momentum, we require that the energy of the entire universe be set to zero, as well and the homogeneous degree of the potential to be -2. By making the change in 23 we get such a potential. The classical Newtonian potential has degree -1, and I_{CM} has degree 2, so by dividing by the square root of I_{CM}, we can set the degree to the desired -2. We can then think of $I_{CM}^{-1/2}$ as a time dependant gravitational coefficient. However because making this change allows us to fix I_{CM} as constant, doing so causes the time dependence to vanish.

It is undesirable that we should have to modify such a powerfully predictive theory as Newtons in this way; new physics should be consistent with old. Perhaps scale is not an indiscernible symmetry of reality, but rather central to the notion of time itself.

2.4. **Preview of the problem of time using the triangle space Hamiltonian.** The Hamiltonian in polar coordinate $\vec{w}$-space is an illuminating way to quickly introduce the problem of time in shape dynamics. We will here give a short explanation. Essentially we wish to remove reference to an external time coordinate t, as is standard in physics, and instead use some internal observable to label events. The fast and lose explanation is that we can relax the scale invariance constraint, which we argued would have set the dilatation momentum to zero for all time and precluded the possibility of Newton's gravity from being complete, and rather use the (no longer restricted) *dilatational momentum as a time coordinate.* To see why this even makes sense, look at a plot in Fig. 3 of the dilitational momentum of a n-body system with zero translational and angular momentum $\vec{L} = \vec{P} = 0$ as well as non-negative total energy $E \geq 0$.

Whenever we impose these restrictions on an N-body system, the dilitational momentum will look like this plot, specifically, it will be an everywhere monotonic increasing function of time[4]. This means $D(t)$ is a diffeomorphism of the line (it preserves the lines essential property, its ordering). This means that we can perform a canonical transformation and use D to label events instead of t, and lose no physically essential information, despite no longer referencing an external clock t.

Let us look at some of the math. First well explicitly state our coordinates:

$$w_1 = R\sin\theta\cos\phi, \quad w_2 = R\sin\theta\sin\phi, \quad w_3 = R\cos\theta \tag{25}$$

Then we can write the Hamiltonian[4]

$$\mathcal{H} = \frac{p_\theta^2 + \sin^{-2}\theta p_\phi^2 + \frac{1}{4}D^2}{2R} + \sqrt{R}U(\theta,\phi) \tag{26}$$

where p_θ, p_ϕ are the momenta conjugate to θ, ϕ. Note that U is independent of $R = I_{CM}$, and $\sqrt{R}U$ is equal to the classical Newtonian potential. The classical Newtonian potential U_{new} written in the polar coordinates of $\vec{w'}$-space, is

$$U_{new} = -\sum_{a<b} \frac{(m_a m_b)^{\frac{3}{2}}(m_a + m_b)^{\frac{-1}{2}}}{\sqrt{R - w_1\cos\phi_{ab} - w_2\sin\phi_{ab}}} \tag{27}$$

with the ϕ angles denoting the binary collision between particles a and b. They are defined as follows according to [4]

$$\phi_{12} = \pi \tag{28}$$

$$\phi_{23} = \arctan\left(2\frac{\sqrt{m_1m_2m_3(m_1+m_2+m_3)}}{m_2(m_1+m_2+m_3) - m_1m_3}\right) \tag{29}$$

$$\phi_{13} = \arctan\left(2\frac{\sqrt{m_1m_2m_3(m_1+m_2+m_3)}}{m_1(m_1+m_2+m_3) - m_2m_3}\right) \tag{30}$$

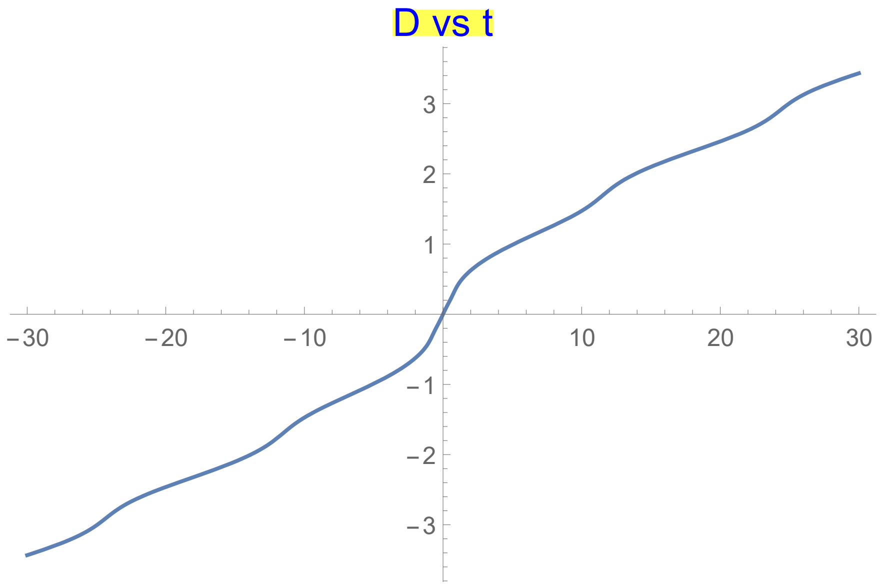


FIGURE 3. Diliational momentum D vs. classical time t for an N-body Newtonian 3-body system.

The Hamiltonian defined in this way is time independent, and $D(t)$ is a diffeomorphism of t, so we can perform a canonical transformation and use the dilatational momentum to label events without loss of any information. Note that this Hamiltonian only has 3 coordinates, R, θ, ϕ, and thier conjugate momenta, while the classical 3-body problem has 9 (the x, y, z coordinates of the 3 particles). This new system has forgotten center of mass and angular state information, losing 6 degrees of freedom. Using the dilatational momentum as time coordinate means there are only 2 non-time coordinates and momenta; the scale and associated momentum are now used to label events like time. The system is now "timeless" because there is no external coordinate along which events proceed, time as dilitational momentum is a directly observable property of the system.

3. General mathematical framework

3.1. **A quick overview.** The program of shape dynamics is to realise a given dynamical system, say a Hamiltonian system, whose generalized coordinates and momenta are additionally restricted to a principle Sim_3-bundle's horizontal subspace, where Sim_3 denotes a group of angle preserving, or *similarity*, transformations of the ambient $\mathbb{R}^3$ spacial coordinates. In other words, it is a *gauge theory* with a global *gauge group* of spacial similarity transformations.

Before we can elucidate our point about shape reductions, we will need a convenient pictorial description of classical mechanics.

3.2. **A picture of classical mechanics.** The diagram in Figure 5 is based on a diagram from Farantos[2], and will help explain the key details of SD succinctly. Nothing about this diagram is related to SD directly; it is merely a way of thinking about classical mechanics in the 'big picture'. To understand this picture and subsequent arguments, an understanding of the concept of *fiber bundle* is necessary. Category theoretic notions will be used. Of particular interest for Lagrangian mechanics is the category of smooth manifolds, called **Diff**, and the *tangent bundle*, a particular type of fiber bundle where the fibers are tangent vector spaces. For additional details see any text on topological manifolds, [9] was suggested to me and I think treats it well. The duel concept of cotangent bundle formed from the associated duel tangent covector spaces are necessary for Hamiltonian mechanics. The weights of this duel transformation (Legendre transform) are the physical masses. A convenient way to understand T is as a functor,

$$T : \mathbf{Diff} \to \mathbf{Vect}(\mathbf{Diff}) \tag{31}$$

from **Diff** to the category of vector bundles over **Diff**, labeled **Vect**(**Diff**). This is in some sense a subcategory of **Diff**, since $\mathbb{R}$-vector spaces are themselves smooth manifolds. The gist of all of this is that for any given manifold $M \in \mathbf{Diff}$, we can always form this new object, in a canonical way, called TM, the tangent bundle of M. It is always another manifold with twice the dimension of the original, and always comes equipped with a projection $\pi_M : TM \to M$, which forgets the attached vector space structure at each point, and gives TM the character of a fiber bundle. T also assigns an associated object to a given smooth map (morphism) in the category $\mathbf{Diff}$), say f, called the differential df (morphism on $\mathbf{Vect}(\mathbf{Diff})$). This rule for mapping objects and morphisms in $\mathbf{Diff}$ to objects and morphisms in $\mathbf{Vect}(\mathbf{Diff})$, among a few other properties, is what makes T a functor.

$$\begin{tikzcd}
 & & T(T^*Q) \arrow[dll, "\pi_{TQ}"'] \arrow[drr, "\pi_{T^*Q}"] & & \\
TQ \arrow[rrrr, "L"] \arrow[d, "\mathcal{L}"] \arrow[drr, "\pi_Q"] & & & & T^*Q \arrow[d, "\mathcal{H}"] \arrow[dll, "\pi_Q^*"] \\
\mathbb{R} & & Q \arrow[d, "U"] & & \mathbb{R} \\
 & & \mathbb{R} & &
\end{tikzcd}$$

FIGURE 4. The picture of classical mechanics

T^*Q then forms the cotangent bundle, or *phase space*, which is the natural realm for discussing Hamiltonian mechanics. A physical theory is formed by: 1. specifying a coordinate manifold Q with (local) coordinates (x^i), 2. a potential function $U : Q \to \mathbb{R}$, 3. the weights g_{ij} (masses) of the Legendre transform L relating tangent vectors (v^i) to tangent covector $(p_i = g_{ij}v^j)$, which combine with each other to define the kinetic metric $2K := v^i p_i = v^i g_{ij} v^j$. We then get a Lagrangian $\mathcal{L} : TQ \to \mathbb{R}$, where $\mathcal{L}(x^i, v^i) = \frac{1}{2} v^i g_{ij} v^j - U(x^i)$ which together with action principles gives the Euler-Lagrange equations dictating the deterministic evolution of the system starting from specified initial conditions $\{x^i(0), v^i(0)\}$. Associated to any Lagrangian one can define the Hamiltonian $\mathcal{H} : T^*Q \to \mathbb{R}$, $\mathcal{H} = v^i p_i - \mathcal{L}$, via the Legendre transform L, and with the same action principles, derive Hamilton's equations of motion dictating evolution. Hamilton's equations are a collection of vector PDE's defined on the tangent bundle of phase space $T(T^*Q)$.

Lets say we start with a Newtonian N-body gravitational system, for simplicity, with all equal masses $m_i = 1$, and gravitational constant $G = 1$. This systems dynamics is governed by either the Lagrangian,

$$\mathcal{L}_N := \sum_{i=1}^{N} \frac{\vec{v}_i \cdot \vec{v}_i}{2} + \sum_{i,j<i} \frac{1}{r_{ij}} \tag{32}$$

with the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}_N}{\partial \vec{v}_i}\right) = \frac{\partial \mathcal{L}_N}{\partial \vec{x}_i} \tag{33}$$

or the Hamiltonian

$$\mathcal{H}_N := \sum_{i=1}^{N} \frac{\vec{p}_i \cdot \vec{p}_i}{2} - \sum_{i,j<i} \frac{1}{r_{ij}} \tag{34}$$

together with Hamilton's equations

$$\frac{\partial \vec{p}_i}{\partial t} = -\frac{\partial \mathcal{H}}{\partial \vec{x}_i} \tag{35}$$

$$\frac{\partial \vec{x}_i}{\partial t} = +\frac{\partial \mathcal{H}}{\partial \vec{p}_i} \tag{36}$$

where again $r_{ij} = |\vec{x}_i - \vec{x}_j|$. Both of these completely determines a solution trajectory $\{\vec{x}_i(t), \vec{p}_i(t)\}$ for a given initial value problem $(\vec{x}_i(0), \vec{p}_i(0)) \in T^*\mathbb{R}^{3N} \simeq \mathbb{R}^{6N}$. (Momentum and velocity coincide in this case since all the $m_i = 1$)

3.3. **Noether's theorem.** Noether's theorem is a key tool in the physicists toolbox because it describes *conservation laws*. Informally it says that if a given Lagrangian has a symmetry, then there is a corresponding conservation law. The quintessential examples are that translational symmetry implies conservation of momentum, rotational symmetry implies conservation of angular momentum, and that time translation symmetry implies conservation of energy. Suppose we have a Lagrangian which does not depend on a generalized coordinate q. Then, via the Euler-Lagrange equations

$$\dot{p} = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) = \frac{\partial \mathcal{L}}{\partial q} = 0 \tag{37}$$

so the generalized momentum p associated to q is conserved. A central point in shape dynamics is a much stronger form of this. We might say a coordinate q physically exists but is dynamically irrelevant. On the surface of the earth, the gravitational potential only depends on changes in height, not changes parallel to the ground. Ignoring friction, Noethers theorem implies that momentum in these directions is conserved. Simply, the horizontal velocity of a thrown ball does not change in time. However in shape dynamics, we are asserting there are redundant additional generalized coordinates in our physical theories that *do not exist at all*. This has the effect of setting the associated momentum to that coordinate *to zero*. It must be zero for all time, and hence is in some rather silly sense, conserved. (if $p = 0$, $\dot{p} = 0$). Suppose there were some hypothetical fourth spacial dimension. If a Lagrangian was extended to this additional space, but did not depend on it, the momentum associated to this new dimension would be constant. But if this new dimension doesn't exist at all, the momentum associated with it would have to be zero. This means a physical Lagrangian should not depend on any physically meaningless coordinates, lest there be no conservation law, which would give the system a changing, and thus non-zero momentum associated to that coordinate. Looking at the standard Lagrangian of Newtonian gravity, we see that center of mass momentum and total angular momentum must be conserved.

$$\mathcal{L} = \sum_i \frac{m_i}{2} v_i^2 + \sum_{i \le j} \frac{m_i m_j}{r_{ij}} \tag{38}$$

where again $r_{ij} = |\vec{x}_i - \vec{x}_j|$, is invariant under translations and rotations

$$\mathcal{L}(\vec{x}_i, \vec{v}_i) = \mathcal{L}(\mathbf{R}\vec{x}_i + \vec{b}, \vec{v}_i) \tag{39}$$

3.4. **The shape gauge group.** Now consider the group of conformal transformations of 3-space given by all transformations of the form

$$\vec{y}_i = s\boldsymbol{R}\vec{x}_i + \vec{b} \tag{40}$$

$$s \in \mathbb{R}^+, \boldsymbol{R} \in SO_3, \vec{b} \in \mathbb{R}^3 \tag{41}$$

a quick note on notation, scalars (elements of $\mathbb{R}$) are denoted by lowercase letters, matrix operators by bold capital letters, in this case elements of SO_3, and 3-vectors with arrow hats. The indices are taken to be in some particle label set, and the transformation should be understood to occur on every particle at once, i.e. the transformation occurs for each i. These transformations can be represented as a subgroup of $GL(4, \mathbb{R})$ with an augmented matrix in the following way.

$$\begin{bmatrix} \vec{y}_i \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} & s\boldsymbol{R} & & \vec{b} \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \vec{x}_i \\ 1 \end{bmatrix} \tag{42}$$

The group axioms are immediately satisfied from this representation[2]. We will call this group Sim_3, the similarity transformations of 3-space.

3.5. **The shape reduction.** In N-body Shape Dynamics the gauge is always global, i.e. there is no spatially localized gauge freedom. Formally we view the configuration manifold of Newtonian gravity $Q = \mathbb{R}^{3N}$ as a principle Sim_3-bundle, which is a fiber bundle who's fibers are Sim_3-torsors. We have a Sim_3 action on the coordinate manifold Q shown above, where $(\vec{x}_i) \in Q$, and projection π to the base space of shape coordinates Q_s,

$$\pi_s : Q \to Q_s \tag{43}$$

[2]this way of representing the group was shown to me by Eric Brussel.

Which annihilates the fibers of the bundle. The fibers are Sim_3-torsors, and the projection is compatible with the group action on the manifold.

$$\forall S \in Sim_3, \forall p \in Q, \tag{44}$$

$$\pi_s(p) = \pi_s(S(p)) \tag{45}$$

The differential of this map, which exists since π is a smooth map in **Diff** and is constructed by the functor T, acts on the tangent bundle,

$$d\pi_s \colon TQ \to TQ_s \tag{46}$$

We have another diagram,

$$\begin{tikzcd} TQ \arrow[r, "d\pi_s"] \arrow[d, "\pi_Q"] & TQ_s \arrow[d, "\pi_{Q_s}"] \\ Q \arrow[r, "\pi_s"] & Q_s \end{tikzcd}$$

Figure 5. Conformal invariant projection and its differential

The kernel of $d\pi_s$ is called the *vertical tangent bundle*, $V := \ker(d\pi_s)$, while its image is the *horizontal tangent bundle*, $H = TQ_s := \text{Image}(d\pi_s)$. This splits the tangent spaces at every point $p \in Q$, $T_pQ = V_p \bigoplus H_p$. This projection can be thought of as purging the system of unphysical (in the pure relational sense) data, i.e. the base space Q_s contains all of the data *internal* to a given configuration, whereas the original space still has certain additional information, specifically an overall scale, translational and angular state. Note that any tangent vector in the original space, or indeed any section of the tangent bundle, will have its *vertical component* destroyed by the projection. Since velocity is a tangent vector, we lose (set to zero) 7 characteristic velocities in the transition to Q_s. These velocities are the translational velocity or *tangent vector* $v_{CM} \in T_{x_{CM}}Q$ of the center of mass or *generalized coordinate* of the universe, coming from the translation ($\vec{b} \in \mathbb{R}^3$) symmetry of Sim_3, and total angular and dilatational velocities, which again are elements of the tangent space at the center of mass, coming from the rotational ($\boldsymbol{R} \in SO_3$) and scaling ($s \in \mathbb{R}^+$) symmetry respectively. Setting each of these velocities to zero additionally sets their associated momentum to zero as well, equivalently, annihilating the vertical tangent bundle also annihilates the vertical cotangent bundle of the phase space.This is because the mass of the system, or inertia tensor[3] in the case of rotation, is a vector space homomorphism from the tangent bundle to the cotangent bundle and must then map zero to zero.

The diagram in fig. 5 acts on the one in fig. 4, dragging it into a a new picture of classical mechanics with pure relational data only. This projection enforces all the 'new rules', specifically about which momenta must be zero, that have been discussed in this paper. Crucially this formulation defers reference to a particular coordinate manifold, and thus imposes *generalized features* which apply to all shape invariant theories on *any manifold*.

4. Arithmetic convenience

Understanding the structure of the gauge group of shape dynamics is central to the entire program. Indeed, the predictions of the entire theory depend on the choice of gauge; for example: *is global scaling symmetry necessary?*, i.e. should the scalar multiplication component of the group be considered or not? The fact that global scaling would imply that Newton's $1/r$ potential could not be the case seems to imply that scaling should not be considered. However I will simply leave some notes about the gauge group chosen in the current literature, Sim_3.

I here build some other modes of representing this group in hoping that they may build some additional understanding about how these groups operate. I will start by using the arithmetic of the complex numbers, as they contain both simple arithmetic, and structure which is quite similar to that of general similarity transformations. I then build an isomorphic representation of Sim_3 using quaternions, following an analogy with the complex construction.

[3]it is important to note that the inertia tensor is the arithmetic representation of the homomorphism from the tangent space to the cotangent space

4.1. **Plane universes and complex numbers.** To build up to an alternative representation of the shape gauge group, we proceed by discussing again triangles. First consider 3 points in a plane, mathematically, 3 complex numbers, $a, b, c \in \mathbb{C}$. The complex numbers *are* a plane, just like the real numbers *are* a line, they just also have some convenient arithmetic structure. We say two triangles $\{a_1, b_1, c_1\}, \{a_2, b_2, c_2\}$ are similar if there exists unique complex numbers $\alpha, \beta \in \mathbb{C}$ such that

$$\begin{aligned} a_1 &= \alpha a_2 + \beta \\ b_1 &= \alpha b_2 + \beta \\ c_1 &= \alpha c_2 + \beta \end{aligned} \tag{47}$$

First note the similarity between this equation and that shown in Eqn. 41. There is a geometrically intuitive explanation of what is going on. Think of the a, b, c objects as points in the plane, and α, β as actions on that plane. α is a scaling and rotation, since if we write in polar form $\alpha = se^{i\theta}$, then α has the effect of rotating $\mathbb{C}$ by θ about the origin, and magnifying by a factor s, again about the origin. Then $\beta = a + bi$ is just plain old translation, by a in the horizontal 1 direction, and by b in the vertical i direction. Both of these actions preserve angles, and it can be shown that these are all of the angle preserving transformations (fixing the point at infinity, otherwise this would be the Mobiüs transformations), or similarity transformations of $\mathbb{C}$. To say that there are complex number α and β is to say that there is a similarity transformation taking the points a_1, b_1, c_1 to a_2, b_2, c_2 preserving angles and therefore they must be similar triangles.

We can now form an equivalence class of triangles in $\mathbb{C}$ by saying two given triangles are equivalent if one of these similarity transformations exists, then we take each of these equivalence classes and treat them as points in a space. This space is the shape space of triangles, its points represent the shapes of possible triangles. This is essentially the process described for the rest of SD, but we started with only 2 dimensional coordinates, given by points in $\mathbb{C}$. If only there were a similar form for a 3D space...

4.2. **The quaternion representation.** Another possibly more arithmetically convenient representation of shape gauge group Sim_3 is as the $\mathrm{Im}(\mathbb{H})$ stable linear functions over the division algebra of quaternions $\mathbb{H}$. This sections follows the notation of [6]. For this representation we send 3-vectors to pure imaginary quaternions $\mathbb{R}^3 \ni \vec{x}_i = x\hat{e}_1 + y\hat{e}_2 + z\hat{e}_3 \to xi + yj + zk = x_i \in \mathrm{Im}(\mathbb{H})$ so that the following transformations are precisely the similarity transformations of $\mathrm{Im}(\mathbb{H}) \simeq \mathbb{R}^3$ (they are isomorphic as vector spaces)

$$Sim_3 \simeq \{f : \mathbb{H} \to \mathbb{H} | f(X) = hX\bar{h} + b, h \in \mathbb{H}^\times, b \in \mathrm{Im}(\mathbb{H})\} \tag{48}$$

This representation bear one internal redundancy, for any $h \in \mathbb{H}^\times$, $-h$ also results in the exact same transformation. This is related to the three sphere being a double cover of $SO(3)$

4.3. **Arithematic and geometry of the quaternions.** Hamilton's quaternions are a 4 dimensional division algebra of $\mathbb{R}$[6]. They obey all the axioms of a field save that multiplication is not commutative. The multiplication is completely determined by a simple equality on 4 basis vectors $1, i, j, k$.

$$i^2 = j^2 = k^2 = ijk = -1 \tag{49}$$

4.4. **Rotations and Dilations.** The property we care about is that quaternions are particularly adept at representing rotations of 3 space, as well as the fourth dimension conveniently also containing dilations. The subspace $\mathrm{Im}(\mathbb{H})$ is precisely the Lie algebra of rotations $\mathfrak{so}(3) \simeq \mathfrak{su}(2) \simeq \mathrm{Im}(\mathbb{H})$. Here is the trick. Suppose we want to rotate about an axis $\hat{\theta}$, by angle θ, and scale by a factor s, fixing the origin. Set $\vec{\theta} = \theta\hat{\theta} = \theta_x i + \theta_y j + \theta_z k \in \mathrm{Im}(\mathbb{H})$. Now

$$h = \exp\left(\frac{\ln s + \vec{\theta}}{2}\right) = \sqrt{s}\exp\frac{\vec{\theta}}{2} \tag{50}$$

$$\bar{h} = \exp\left(\frac{\ln s - \vec{\theta}}{2}\right) = \sqrt{s}\exp\frac{-\vec{\theta}}{2} \tag{51}$$

$$Y = hX\bar{h}, X, Y \in \mathrm{Im}(\mathbb{H}) \tag{52}$$

rotates and scales any vector X to Y. In this way the quaternions are the Lie algebra of rotations and scaling of 3 space, though scalings commute with rotations so their Lie algebra structure is trivial. However it is interesting that the real part of the quaternions, $\mathrm{Re}(\mathbb{H})$ represent the scaling, and the imaginary part the rotations. This explains how the representation in Eqn. 48 contains the rotational and dialational information of the group Sim_3.Translations are captures by the additive structure of the underlying vector space of the quaternion algebra.

4.5. **Multiplicative Structure.** There is a 'neat trick' available for calculating the angular and dilatational momentum for a system in a given state $\vec{x}_i, \vec{p}_i$. If we have two vectors $x, p \in \mathrm{Im}(\mathbb{H})$, then

$$xp = -\vec{x} \cdot \vec{p} + \vec{x} \times \vec{p} \cdot (i, j, k) \tag{53}$$

If we represent our state as pure imaginary quaternions $x_i, p_i \in \mathrm{Im}(\mathbb{H})$, with center of mass as the origin, then

$$\sum_i x_i p_i = -D + L \tag{54}$$

where $D \in \mathrm{Re}(\mathbb{H})$ is the dialatational momentum, and $L \in \mathrm{Im}(\mathbb{H})$ is the angular momentum. Understanding the restriction

$$\sum_i x_i p_i = 0 \tag{55}$$

in this context may be illuminating to the program.

5. Conclusion

Shape Dynamics is a gauge theory with a gauge group of angle preserving transformations acting on the coordinate manifold of configurations, endowing this manifold with the structure of a principle bundle. It reduces the available dynamics of a given physical theory by asserting that it must evolve through 'horizontal' changes only, which is called 'best matching' in the N-body case. By asserting that only relational data exists, and thus that the gauge group is physically irrelevant and non-existent, we restrict to the base space of the principle bundle, and enforce horizontality, forcing 7 generalized momenta to 0 for all time. All of this can be seen in a concise pictures, a projection dragging a description of general classical systems into a shape dynamic framework. This generality means that it ought to easily extend to more relevant contexts and general 3-manifolds which are consistent with observations about the global topology of the observable universe. The logic attached to the diagrams shown do not care about which manifold I 'plug in' for Q, the results derived still apply for *any* smooth manifold. All that is required is specification of a Sim_3 action.

It is already known that Shape Dynamics applied to general manifolds leads to a pseudo-equivalent formulation of general relativity, where predictions of each theory coincide in many cases and not others. Having a concise explanation of how this comes about, following similar diagrammatic form could help elucidate these distinctions and the logical arguments involved in a more digestible manner.

Another key point that was not discussed in the paper is the elimination of time by use of a time reparametrization invariant action, called the Jacobi action.

Combining the spacial restrictions with time reparametrization invariance leads to Geometrodynamics, which is a shape gauge invariant theory predicting the same results as GR in many, but not all, contexts.

References

[1] Julian Barbour. Shape dynamics. an introduction, 2011.
[2] Stavros C. Farantos. *The Geometry of Hamiltonian Mechanics*, pages 13–31. Springer International Publishing, Cham, 2014.
[3] Flavio Mercati Julian Barbour, Tim Koslowski. The arrow of time, complexity and bulk entropy in gravity.
[4] Flavio Mercati Julian Barbour, Tim Koslowski. A gravitational origin of the arrows of time. 2013.
[5] Flavio Mercati Julian Barbour, Tim Koslowski. Identification of a gravitational arrow of time. *Physical Review Letters*, 2014.
[6] R. Remmert M. Koecher. Hamilton's quaternions. In J. H. Ewing, editor, *Numbers*, chapter 7, pages 189–220. Springer-Verlag, New York, 1990.
[7] Ernst Mach. *The Science of Mechanics.* The Open Court Publishing Co., 1919.
[8] Flavio Mercati. A shape dynamics tutorial, 2014.
[9] John W. Milnor. *Topology from the Differentiable Viewpoint.* Princeton University Press, 1997.
[10] Henri Poincarè. *Science and Hypothesis.* Walter Scott, 1905.